UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/761,451	01/20/2004	Dong Yu	M61.12-0582	3046

27366 7590 06/08/2009

WESTMAN CHAMPLIN (MICROSOFT CORPORATION)
SUITE 1400
900 SECOND AVENUE SOUTH
MINNEAPOLIS, MN 55402

EXAMINER	
SHAH, PARAS D	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
06/08/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/761,451	YU ET AL.	
	Examiner	Art Unit	
	PARAS SHAH	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *30 March 2009*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,3,4,7,15-19 and 26-36* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,3,4,7,15-19 and 26-33, 35, 36* is/are rejected.

7)☐ Claim(s) *34* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This communication is in response to the Amendments and Arguments filed on 03/30/2009. Claims1, 3, 4, 6, 7, 14-19, and 21-23 remain pending and have been examined, while claims 6, 14, and 21-23 have been cancelled, and claims 26-36 have been newly added. The Applicants' amendment and remarks have been carefully considered, but they do not place the claims in condition for allowance.

2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

Response to Amendments and Arguments

3. Applicant's arguments (pages 7-9) filed on .03/30/2009 with regard to claims 1 and 7 have been fully considered but they are moot in view of new grounds for rejection.

With regards to claim 26, which incorporates subject matter from claims 22 and 23, the Applicants argue that Hoffman does not teach or suggest the temporary adding words to a user lexicon sine Hoffman discards words according to a FIFO principle. The Examiner respectfully disagrees with this assertion. In Hoffman, paragraphs [0015] and [0031], Hoffman teaches the substitution of words based on two criteria. The criteria includes the frequency and oldest date of use. Hence, the words that are added are temporarily stored based on how often the user uses such words and when it was last used. Thus, each word contained in the lexicon is temporary with respect to the criteria that Hoffman defines.

Information Disclosure Statement

4. The information disclosure statement filed 03/30/2009 fails to comply with the provisions of 37 CFR 1.97, 1.98 and MPEP § 609 because it does not provide a translated copy of the Chinese Office Action and also no copy of the Chinese Office Action has been supplied. It has been placed in the application file, but the information referred to therein has not been considered as to the merits. Applicant is advised that the date of any re-submission of any item of information contained in this information disclosure statement or the submission of any missing element(s) will be the date of submission for purposes of determining compliance with the requirements based on the time of filing the statement, including all certification requirements for statements under 37 CFR 1.97(e). See MPEP § 609.05(a).

Claim Rejections - 35 USC § 112

5. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

6. Claims 1 and 7 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The newly added limitation of "… to selectively change at least one HMM parameter with an existing pronunciation" is not

supported by the specification. The cited sections mentioned by the Applicant, in the Applicant's Remarks, page 7, lines 16-25 describes the use of HMM and transition probability. The cited portions are described in the Background Section and describes the modeling of a word. The second section denoted by the Applicant only states that a "probability of newly observed known probabilities might also be increased." This portion does not provide support since the section does not state that such probability that is increased occurs via a HMM parameter or an Acoustic model. The section merely describes that the language model is updated in lines 11 of page 21. There is no mention in any of the mentioned pages or anywhere else in the Specification that such change is a change to a HMM parameter. Hence, the Applicant's newly amended limitations are not supported by the Specification, where it would reasonably convey to one skilled in the art that the Applicant had possession of the invention.

7. Claims 1 and 7 are rejected under 35 U.S.C. 112, first paragraph, because the specification, while being enabling for the correction of speech and updating of a language model, does not reasonably provide enablement for "change at least one HMM parameter associated with an existing pronunciation". The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to use the invention commensurate in scope with these claims. The limitation of "change at least one HMM parameter associated with an existing pronunciation" for which the Applicant is claiming to be their invention, does not provide enablement as page 21, lines 21-24 where the probability of a newly observed pronunciation is increased. It does not describe to one skilled in the art as to how this

increase takes place, what is increasing and by how much is it increasing. Further, the Background section does not provide enablement for the mentioned limitation since it merely describes how modeling of a word occurs using HMM and transition probabilities. The Applicant has failed to link the background section with Applicant's change of an HMM parameter since the Applicant has not provided adequate description that would enable one of ordinary skilled in the art to make and use the invention, specifically the changing in a HMM parameter.

8. Claims 3, 4, 15-19 are rejected for being dependent upon a rejected base claim.

Claim Rejections - 35 USC § 101

9. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claim 26, 35, and 36 are rejected under 35 U.S.C. 101 as not falling within one of the four statutory categories of invention. Supreme Court precedent[1] and recent Federal Circuit decisions[2] indicate that a statutory "process" under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. While the instant claim(s) recite a series of steps or acts to be performed, the claim(s) neither transform underlying subject matter nor positively tie to another

---

[1] *Diamond v. Diehr*, 450 U.S. 175, 184 (1981); *Parker v. Flook*, 437 U.S. 584, 588 n.9 (1978); *Gottschalk v. Benson*, 409 U.S. 63, 70 (1972); *Cochrane v. Deener*, 94 U.S. 780, 787-88 (1876).

[2] *In re Bilski*, 88 USPQ2d 1385 (Fed. Cir. 2008).

statutory category that accomplishes the claimed method steps, and therefore do not qualify as a statutory process. For example the learning of pronunciation method including steps of receiving, analyzing and responding is of sufficient breadth that it would be reasonably interpreted as a series of steps completely performed mentally, verbally or without a machine. The Applicant has provided no explicit and deliberate definitions of "detecting", “inferring,” "selectively learning," or “adding” to limit the steps to the electronic form of the” email question," and the claim language itself is sufficiently broad to read on a human listening to another individual speaking. The human sees the other individual has made a correction to the words transcribed by the human. Determining done by the human as to how close the transcribed text differs from the text changed by the other individual by looking at the character differences. The human then adds the corrected word on to a piece of paper comprising words spoken by the individual. The human can edit the list (pruning) based on a timing threshold and a frequency of occurrence.

Claim Rejections - 35 USC § 103

10. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

11. Claims 1, 3, 4, 27-31 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stevens *et al.* (US 6,9112,498) in view of Honda *et al.* (US 6,879,956).

As to claim 1, Stevens *et al.* teaches

a computer-implemented speech recognition system comprising:

a microphone to receive user speech (see col. 26, line 30, microphone)

a speech recognition engine coupled to the microphone (see col. 26, lines 27-29, speech recognition system and lines 31-32, speech recognition performed based on input) and being adapted to recognize the user speech and provide a textual output on a user interface (see col. 26, lines 31-39, where the user can edit dictated text and col. 27, lines 26--33 and see Figure 11 A, results are displayed to the user and allows user correction); and

wherein the system is adapted to recognize a user changing the textual output and automatically (see col. 26, lines 17-24, and lines 31-39, where the system recognizes an edit or revision), selectively adapt the speech recognition engine to learn from the change (see col. 26, lines 31-39, acoustic models are adapted based on user changes and recognition results); and

wherein the recognition engine is configured to determine if the user's pronunciation caused the error, and to selectively change at least one parameter associated with an existing pronunciation (see col. 27, lines 1-21, when the difference is within a threshold then the acoustic model is adapted).

However, Stevens does not specifically teach to selectively change at least one HMM parameter.

Honda does teach to selectively change at least one HMM parameter (see col. 7, lines 6-12, where acoustic model adaptation alter parameters such as the average vale and variance that defines the transition probability of the HMM)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the correction of dictating speech of Stevens *et al.* with the inclusion of to selectively change at least one HMM parameter as taught by Honda. The motivation to have combined the references involves the prevention of burdening the user for speech recognition and to adaptively adapt to on-line recognition thereby increasing speech recognition performance (see Honda, col. 2, lines 1-6, lines 21-24).

As to claims 3 and 4, Stevens in view of Honda teaches all of the limitations as in claim 1

Furthermore, Honda teaches wherein the HMM parameter is an output probability (see col. 7, lines 6-12, where the altering defines the output probability or transition probability.

As to claim 27, Stevens in view of Honda teaches all of the limitations as in claim 1.

Furthermore, Stevens teaches wherein the system is configured to adapt the speech recognition engine if a distance between the user's pronunciation and a pronunciation of the changed textual output is below a threshold (see col. 27,

lines 1-21, where a threshold is used to determine whether to adapt an acoustic model).

However, Stevens does not specifically teach the adapting when above a threshold.

It would have been obvious to one of ordinary skilled in the art to have used any type of threshold such as above a threshold, as a design criteria, in order to obtain predictable result of adapting an acoustic model based on user edit in order to successfully recognize the word (see Stevens, col. 26, lines 1-21) upon further occurrences of the word by the user.

As to claim 28, Stevens in view of Honda teaches all of the limitations as in claim 27.

Furthermore, Stevens teaches wherein the threshold is pre-selected (see col. 27, lines 13, threshold employed).

As to claim 29, Stevens in view of Honda teaches all of the limitations as in claim 27.

Furthermore, Stevens teaches wherein the threshold is dynamic (see col. 27, lines 13, tunable threshold employed).

It would have been obvious to one of ordinary skilled in the art to have used a dynamic threshold in order to obtain predictable result of adapting an acoustic model based on user edit in order to successfully recognize the word

(see Stevens, col. 26, lines 1-21) upon further occurrences of the word by the user, where a dynamic threshold can be used based upon how much adaptation for the system should occur based upon the correction or revision made by the user.

As to claim 30, Stevens in view of Honda teaches all of the limitations as in claim 27.

Furthermore, Stevens teaches wherein the system is configured to identify the pronunciation of the changed textual output using a lattice constructed using phoneme sequences in a recognition result (see col. 26, lines 46-67, where the system builds an acoustic model containing a phonetic representation of a word based upon the edit or correction.)

As to claim 31, Stevens in view of Honda teaches all of the limitations as in claim 27.

Furthermore, Stevens teaches wherein the distance is calculated based on an acoustic model score on the pronunciation of the changed textual output (see col. 27, lines –20, where an Acoustic model score based on the change and the original is compared).

12. Claims 7 and 15-19, and 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stevens *et al.* in view of Honda *et al.* in view of Beaufays *et al.* ("Learning Linguistically Valid Pronunciations from Acoustic Data", Sept. 2003).

As to claim 7, Stevens *et al.* teaches a method of learning with an automatic speech recognition system, the method comprising:

detecting a change to dictated text (see col. 26, lines 17-24, and lines 31-39, where the system recognizes an edit or revision);

inferring whether the change is a correction, or editing (see col. 26, lines 17-24, and lines 31-39, where the system recognizes an edit or revision); and

wherein inferring whether the change is a correction includes comparing a speech recognition engine score of the dictated text and of the changed text (see col. 27, lines 1-10, where a score is computer based on the user correction or revision and compares it to an original acoustic model).

if the change is inferred to be a correction, selectively learning from the nature of the correction without additional user interaction (see col. 26, lines 31-39, acoustic models are adapted based on user changes and recognition results).

wherein selectively learning from the nature of correction includes:

determining if a user's pronunciation deviated from an existing pronunciation known by the system by doing a comparison (see col. 27, lines 1-21, where a comparison of an original acoustic model and a edited acoustic model is compared);

determining if the corrected word exists in the user lexicon (see col. 26, lines 52-60, where the best representation of the word is determined for adaptation), selectively changing a parameter associated with the pronunciation (see col. 27, lines 1-21, when the difference is within a threshold then the acoustic model is adapted)

However, Stevens does not specifically teach to selectively change at least one HMM parameter.

Honda does teach to selectively change at least one HMM parameter (see col. 7, lines 6-12, where acoustic model adaptation alter parameters such as the average vale and variance that defines the transition probability of the HMM)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the correction of dictating speech of Stevens *et al.* with the inclusion of to selectively change at least one HMM parameter as taught by Honda. The motivation to have combined the references involves the prevention of burdening the user for speech recognition and to adaptively adapt to on-line recognition thereby increasing speech recognition performance (see Honda, col. 2, lines 1-6, lines 21-24).

However, Stevens in view of Honda do not specifically teach the comparison being done using a forced alignment of a wave based on at least one context word.

Beaufays et al. does teach the forced alignment of a wave (see page 2594, sect. 2, step 1, where a forced alignment of two waveforms is performed)

based on at least one context word (see page 2574, sect. 2, step 2, where the context is use to determine a region of worst acoustic match and in step 3, alternative pronunciations are suggested).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the correction of dictated speech presented by Stevens in view of Honda with the inclusion of alignment between two words as taught by Beaufays. The motivation to have combined the references involves the ability to learn pronunciation from data (see Beaufays, Abstract) in order to reduce which would benefit the teachings of Stevens in view of Honda to lessen the speech recognition errors upon further passes by the user.

As to claim 15, Stevens in view of Honda in view of Beaufays teaches all of the limitations as in claim 14.

Furthermore, Stevens teaches the identification of the corrected word (see col. 26, lines 30-24,, user correction or edit)

Furthermore, Beaufays teaches wherein determining if the user's pronunciation deviated from existing pronunciations includes identifying in the wave the pronunciation (see page 2574, sect. 2, step 2, region of worst acoustic match).

As to claim 16, Stevens in view of Honda in view of Beaufays teaches all of the limitations as in claim 15.

Furthermore, Beaufays teaches wherein building a lattice based upon possible pronunciations of the corrected word and the recognition result. (see page 2594, step 3, suggestion of alternative pronunciations, where alternative phone sequences of are proposed.)

As to claims 17, Stevens in view of Honda in view of Beaufays teaches all of the limitations as in claim 16.

Furthermore, Beaufays teaches wherein generating a confidence score based at least in part upon the distance of the newly identified pronunciation the possible pronunciations (see page 2594, sect. 2, step 4, pronunciation scoring based upon likelihood of alignment of the pronunciations, closeness to each other).

As to claim 18, Stevens in view of Honda in view of Beaufays teaches all of the limitations as in claim 16.

Furthermore, Stevens teaches generating a confidence score based at least in part upon an Acoustic model score of the pronunciation with the possible pronunciations (see col. 27, lines 1-20, where an acoustic model score is determined for the edit and the original acoustic model and then it is compared to a threshold.).

Furthermore, Beaufays teaches the scoring performed on multiple pronunciations (see page 2594, sect. 2, step 4, pronunciation scoring on alternative pronunciations)

As to claim 19, Stevens in view of Honda in view of Beaufays teaches all of the limitations as in claim 17.

Stevens teaches wherein selectively learning the pronunciation includes comparing the confidence score to a threshold (see col. 27, lines 13-21, where a threshold is used).

As to claim 33, Stevens in view of Honda in view of Beaufays teaches all of the limitations as in claim 17.

Furthermore, Stevens does teach the use of confidence score (see col. 27, lines 1-20, where Stevens uses an acoustic model scores and a distance measure between them to determine adaptation of acoustic model.)

However, Stevens in view of Honda in view of Beaufays do not specifically teach the confidence score calculated using the function: $1-(1-p(d, AM))^f$; where p(d, AM) is the probability that a pronunciation with a distance d and AM score is the correct pronunciation, and f is the frequency that the same recognized pronunciation is pronounced.

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have used such a function, which is well known in

statistics where the function describes the probability that the pronunciation is not correct, and the subtraction by 1 denotes that the probability that the pronunciation is correct, where the f is the number of trials that have been performed for determining the incorrectness of the pronunciation upon various trials. Such a function as described would have been obvious in order to obtain the predictable result of obtaining a score, which results in a particular level of confidence, in determining how reliable the pronunciation and distance measures are for adapting a speech recognition system as taught in Stevens (See col. 27, lines 1-20). Such a function is similar to a binomial distribution, which is based on the number of trials and is of the form $Pr(K=k)=(n\ k)p^k\ (1-p)^{n-k}$. The binomial distribution theorem allows for the probability of determining unsuccessful outcomes, where if k successes is zero, and n is the number of trials, such an equation would reduce to $(1-p)^n$, which denotes the probability of not getting successful outcome in n trials. The subtraction of one as claimed yields the probability of obtaining the successful or correct pronunciation

.

13. Claims 26, 35, and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stevens in view of Deligne (US 7,409,345) in view of Hoffman *et al.*(US 2003/0139922).

As to claim 26, Stevens *et al.* teaches a method of learning with an automatic speech recognition system, the method comprising:

detecting a change to dictated text (see col. 26, lines 17-24, and lines 31-39, where the system recognizes an edit or revision);

inferring whether the change is a correction, or editing (see col. 26, lines 17-24, and lines 31-39, where the system recognizes an edit or revision); and

wherein inferring whether the change is a correction includes comparing a speech recognition engine score of the dictated text and of the changed text (see col. 27, lines 1-10, where a score is computer based on the user correction or revision and compares it to an original acoustic model).

if the change is inferred to be a correction, selectively learning from the nature of the correction without additional user interaction (see col. 26, lines 31-39, acoustic models are adapted based on user changes and recognition results).

wherein selectively learning from the nature of correction includes selectively adding at least one word pair to the user's lexicon (see col. 26, lines 52-53, acoustic models are built based upon the correction or revision of the user and see col. 21, lines 58-64, were the system adds words to a vocabulary based upon misrecognized words) (e.g. The generation of an acoustic model based on the correction is an alternative model for an existing word constituting a word-pair).

However, Stevens does not specifically teach the adding of words to a lexicon based on pronunciation variation.

Deligne does teach the addition of words to a lexicon based on pronunciation variants of a word constituting a word pair (see col. 3, lines 35-52, where words are added to a lexicon based on variation).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the learning of pronunciation as taught by Stevens with the addition of words to a lexicon as taught by Deligne. The motivation to have combined the references involves the ability to recognize variations of pronunciation made by the user in order to improve speech recognition accuracy (see Deligne col. 3, lines 35-40 and col. 2, limes 17-20).

However, Stevens in view of Deligne does not specifically teach the temporarily addition of a word to a lexicon and wherein the length of time the word pair is added to the user's lexicon is based at least partially upon the most recent time the word pair is observed and the relative frequency that the pair has been observed in the past.

Hoffmann et al. teaches the addition of a word to a lexicon (vocabulary) is based at least partially upon the most recent time the word pair is observed (see [0015], FIFO, where the words not used for a long time are omitted) and the relative frequency (see [0015] and [0031], frequency of occurrence, that the pair has been observed in the past.) (e.g. In the previous response the Applicant argues that since it is based on FIFO principle the time criteria is not used. Such argument is respectfully traversed since Hoffman uses date information in combination of frequency for the length of time such words exists in the lexicon

before substitution of another word, where if a word is not used by the user then it may be removed)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech recognition system as taught by Stevens in view of Deligne *et al.* with the updating a vocabulary depending on frequency and time as taught by Hoffmann *et al.*. The motivation to have combined the references involves continuous renewal of the vocabulary to eliminate word snot used often and those not used for a long time (See Hoffmann *et al.*, [0015]).

As to claims 35 and 36, Stevens in view of Deligne in view of Hoffman teach all of the limitations as in claim 26, above.

Furthermore, Hoffman teaches wherein one word-pair is added to the user's lexicon temporarily for a specific time period (see [0009], [0015], and [0031], where time information, specifically frequency and oldest date of use is utilized to determined substitution).

However, Stevens in view of Deligne in view of Hoffman do not specifically teach the period being one day or 2 days.

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have used a period of 1 or 2 days based upon number of times the user uses the speech recognition system in order to maintain the user vocabulary at 1,000 words and to remove older word pairs and add new

word pairs that have been recognized as taught by Hoffman (see [0008] and [0031), where the vocabulary would be up to date.

14. Claim 32 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stevens *et al.* in view of Honda *et al.* as applied in claim 17, above and further in view of Thong (US 2003/0187643) in view of Rajput (US 2004/0017180).

As to claim 32, Stevens in view of Honda teach all of the limitations as in claim 17, above.

Furthermore, Stevens does teach the calculation of a distance (see col. 27, lines 1-20).

However, Stevens in view of Honda do not specifically teach the use of a phone confusion matrix and dynamic time warping.

Van Thong teaches the use of a phone confusability matrix being used to compute a distance metric (see [0078], phoneme confusion matrix).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech recognition system as taught by Stevens in view of Honda with the use of a phone confusability matrix as taught by Van Thong in order to compute a distance metric by comparing one phoneme string to another where a phoneme matrix allows the determination of phonetically similar word alternatives and where the use of matrix is used for the distance calculation (see Van Thong [0069]-[0078]) to make the determination of similarity.

However, Stevens in view of Honda in view of Van Thong do not specifically teach the use of a phone confusion matrix and dynamic time warping

Rajput does disclose the use of Dynamic Time warping to align alternative pronunciations with the dictated speech (see [0018]).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech recognition system as taught by Stevens in view of Honda in view of Van Thong with the use of Dynamic Time warping as taught by Rajput in order to determine the similarity of two sequences through the alignment so that a speech recognition system can be adapted to a pronunciation for improving accuracy and speed (see Rajput [0008], and [0018]).

Allowable Subject Matter

15. Claim 34 is objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

16. The following is a statement of reasons for the indication of allowable subject matter: None of the cited reference either alone or in combination thereof teach or suggest the confidence score being calculated using the function "1/[d/f/log(len 1 +len2)], where d is the distance between the recognized pronunciation and a best match in a lexicon, f is a frequency that the same pronunciation is pronounced, and lenl and

len2 are the lengths of phonemes in a new pronunciation and the closest pronunciation, respectively." as recited in claim 34.

Conclusion

17. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Bielby et al. (US 5,644,680) is cited to disclose update of markov models for speech recognition. Heckerman et al. (US 6,263,308) is cited to disclose speech recognition where acoustic models are improved. Lewis (US 6,577,999) is cited to disclose management of acoustic models and user correction of speech. Mangu et al. (SU 2002/0165716) is cited to disclose error correction for speech decoding.

The NPL document by Fosler et al. ("Automatic learning of word pronunciation from data") is cited to disclose learning of word pronunciation based on phone recognition.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PARAS SHAH whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:00a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on (571)272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/David R Hudspeth/
Supervisory Patent Examiner, Art Unit 2626

/Paras Shah/
Examiner, Art Unit 2626

06/04/2009